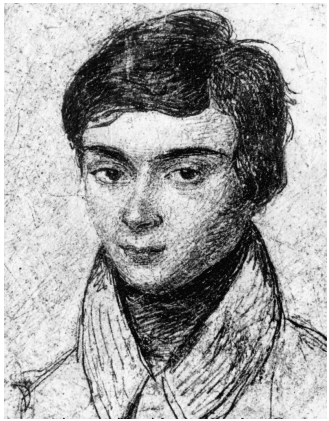
Lecture 14: Finite Fields

MAT205 Abstract Algebra II

Ma, Jia-Jun

Xiamen University Malaysia

Évariste Galois and the Birth of $GF(p^n)$



Évariste Galois (1811–1832)

- Died at 20 in a duel, 31 May 1832.
- 1830 paper *Sur la théorie des nombres* (Férussac's *Bulletin*, tome XIII).
- First appearance of $\mathbb{F}_{p^n}$. The notation $GF(p^n)$ stands for "Galois Field".
- Three Galois theorems we will re-prove:
  - existence of a field of size p^n;
  - $GF(p^n)^\times$ cyclic;
  - $GF(p^d) \subseteq GF(p^n) \iff d \mid n$.

From Galois's 1830 Paper

PURES ET APPLIQUÉES. 399

bre premier donné p, et qu'on cherche, dans cette convention, les solutions d'une équation algébrique $Fx = 0$, ce que M. Gauss désigne par la notation $Fx \equiv 0$, on n'a coutume de considérer que les solutions entières de ces sortes de questions. Ayant été conduit par des recherches particulières à considérer les solutions incommensurables, je suis parvenu à quelques résultats que je crois nouveaux.

Soit une pareille équation ou congruence, $Fx = 0$, et p le module. Supposons d'abord, pour plus de simplicité, que la congruence en question n'admette aucun facteur commensurable, c'est-à-dire qu'on ne puisse pas trouver trois fonctions φx, ψx, χx telles que

$$\varphi x . \psi x = Fx + p\chi x.$$

Dans ce cas, la congruence n'admettra donc aucune racine entière, ni même aucune racine incommensurable de degré inférieur. Il faut donc regarder les racines de cette congruence comme des espèces de symboles imaginaires, puisqu'elles ne satisfont pas aux questions des nombres entiers, symboles dont l'emploi, dans le calcul, sera souvent aussi utile que celui de l'imaginaire $\sqrt{-1}$ dans l'analyse ordinaire.

C'est la classification de ces imaginaires, et leur réduction au plus petit nombre possible, qui va nous occuper.

Appelons i l'une des racines de la congruence $Fx = 0$, que nous supposerons du degré ν.

Considérons l'expression générale

$$\text{(A)} \qquad a + a_1 i + a_2 i^2 + \ldots + a_{\nu-1} i^{\nu-1},$$

où $a, a_1, a_2, \ldots, a_{\nu-1}$ représentent des nombres entiers. En donnant à ces nombres toutes les valeurs, l'expression (A) en acquiert p^ν, qui jouissent, ainsi que je vais le faire voir, des mêmes propriétés que les nombres naturels dans la *théorie des résidus des puissances.*

Ne prenons des expressions (A) que les $p^\nu - 1$ valeurs où $a, a_1, a_2, \ldots, a_{\nu-1}$ ne sont pas toutes nulles : soit α l'une de ces expressions.

Si l'on élève successivement α aux puissances 2^e, $3^e, \ldots$, on aura une suite de quantités de même forme [parce que toute fonction de i peu se réduire au $(\nu - 1)^{ième}$ degré]. Donc on devra avoir $\alpha^n = 1$, n étant un certain nombre; soit n le plus petit nombre qui soit tel que l'on ait

II.4 On the theory of numbers 73

If we wish to have the congruence of least degree on which our primitive root depends, it is necessary to eliminate i between the two equations

$$i^3 = 2, \qquad \alpha = i - i^2.$$

In this way one obtains

$$\alpha^3 - \alpha + 2 = 0.$$

It will be convenient to represent by i the root of this equation, so that

$$i^3 - i + 2 = 0, \qquad (i)$$

and to take it as basis of the imaginaries, so that one will have all the imaginaries in the form

$$a + a_1 i + a_2 i^2$$

by raising i to all powers, and reducing by the equation (i).

The principal benefit of the new theory that we have just expounded is to carry over to congruences the property (so useful in ordinary equations) of admitting precisely as many roots as there are units in the order of their degree.

The method to get all these roots will be very simple. First one can always prepare [adjust] the given congruence $Fx = 0$ in such a way that it does not have equal roots any more, or in other words, so that it will have no factor

in common with $F'x = 0$, and the way to do this is clearly the same as for ordinary equations.

After that, to get the integer solutions it will suffice, as Mr Libri seems to have been the first to notice, to seek the greatest common factor of $Fx = 0$ and $x^{p-1} = 1$.

If one now wishes to have the imaginary solutions of the second degree one will seek the greatest common factor of $Fx = 0$ and $x^{p^2-1} = 1$, and in general the solutions of order ν will be given by the greatest common factor of $Fx = 0$ and $x^{p^\nu-1} = 1$.

It is above all in the theory of permutations, where one forever needs to vary the form of indices, that consideration of the imaginary roots of congruences appears to be indispensable. It gives a simple and easy means of recognising the cases in which a primitive equation is soluble by radicals, of which I shall try to give an idea in two words.

Let $fx = 0$ be an algebraic equation of degree p^ν. Suppose that the p^ν roots are denoted by x_k, where the index k is given the p^ν values determined by the congruence $k^{p^\nu} = k \pmod p$.

Original French — Liouville's *Œuvres* (1846), p. 399. English translation — Neumann (EMS 2011).

Galois treats the roots as "imaginary symbols … as useful as $\sqrt{-1}$", written $a + a_1 i + \cdots + a_{\nu-1} i^{\nu-1}$ — an explicit $GF(p^\nu)$ (e.g. $\mathbb{F}_{27}$ via $i^3 = 2$ over $\mathbb{F}_3$).

Characteristic and Prime-Power Order

- **Definition.** $\operatorname{char} F =$ least $n \geq 1$ with $n \cdot 1_F = 0$; or 0 if no such n.
- **Lemma.** $\operatorname{char} F$ is 0 or prime. *(If $n = ab$ then $(a \cdot 1_F)(b \cdot 1_F) = 0$ in a field — one factor vanishes, contradicting minimality.)*
- **Prime subfield.** The smallest subfield of F: $\mathbb{Q}$ if $\operatorname{char} = 0$ (fraction field of the image of $\mathbb{Z} \to F$), $\mathbb{F}_p$ if $\operatorname{char} = p$ (the image itself).
- **Theorem.** Every finite field has $|F| = p^n$ for some prime p and $n \geq 1$. *(F is a finite-dim $\mathbb{F}_p$-vector space.)*

The Frobenius Homomorphism

Let F be a field of characteristic p.

- **Definition.** The *Frobenius* is $\sigma : F \to F$, $\sigma(a) = a^p$.
- **Theorem.** σ is a ring homomorphism.
- Key step: in $(a+b)^p = \sum \binom{p}{k} a^k b^{p-k}$, the middle binomial coefficients are divisible by p (numerator has factor p, denominator $k!$ does not for $1 \le k \le p-1$). So they vanish in characteristic p, leaving $(a+b)^p = a^p + b^p$.
- **Corollary.** On a finite field, σ is an automorphism (injective $\Rightarrow$ surjective by finiteness).
- Notation: $\sigma^k(a) = a^{p^k}$.

Fixed Subfield of a Group Action

Setup. G acts on a field E by automorphisms ($G \to \mathrm{Aut}(E)$ a group hom).

Definition.

$$E^G := \{\, a \in E : g(a) = a \text{ for all } g \in G \,\}.$$

Theorem. E^G is a subfield of E. (Each g is a ring hom with $g(1) = 1$, so the fixed set is closed under $+, \cdot, ^{-1}, -$.)

The pairing "subgroup $\leftrightarrow$ fixed subfield" is the central object of Galois theory.

Fixed Subfield: Three Examples

- $E = \mathbb{C}$, $G = \langle\, \overline{\cdot} \,\rangle$ (complex conjugation): $E^G = \mathbb{R}$.
- $E = \mathbb{Q}(\sqrt{2})$, $G = \langle \tau \rangle$, $\tau(\sqrt{2}) = -\sqrt{2}$: $E^G = \mathbb{Q}$.
- $E = \overline{\mathbb{F}_p}$, $G = \langle \sigma^n \rangle$ (Frobenius iterated n times):

$$E^G \;=\; \{\, a \in \overline{\mathbb{F}_p} : a^{p^n} = a \,\}.$$

  This is the finite field $\mathbb{F}_{p^n}$.

Definition: $\mathbb{F}_{p^n}$ Inside $\overline{\mathbb{F}_p}$

Fix a prime p. Let $\overline{\mathbb{F}_p}$ be an algebraic closure of $\mathbb{F}_p$; σ the Frobenius on $\overline{\mathbb{F}_p}$.

Definition. For each $n \geq 1$,

$$\mathbb{F}_{p^n} := \overline{\mathbb{F}_p}^{\langle \sigma^n \rangle} = \{\, a \in \overline{\mathbb{F}_p} : a^{p^n} = a \,\}.$$

- Subfield of $\overline{\mathbb{F}_p}$ *automatically* (fixed-subfield theorem).
- $\mathbb{F}_{p^1} = \mathbb{F}_p$ (Fermat).
- Size $|\mathbb{F}_{p^n}| =?$ — count roots of $x^{p^n} - x$ in $\overline{\mathbb{F}_p}$.

No-Repeated-Roots Criterion

Criterion. For $0 \neq f \in F[x]$:

$$f \text{ has no repeated roots in } \overline{F} \iff \gcd(f, f') = 1 \text{ in } F[x].$$

(f' = formal derivative $\sum a_i x^i \mapsto \sum i\, a_i x^{i-1}$.)

Proof.

- If $(x-\alpha)^2 \mid f$, write $f = (x-\alpha)^2 q$; product rule gives $f' = (x-\alpha)(2q + (x-\alpha)q')$, so $(x-\alpha) \mid \gcd(f, f')$.
- Conversely, if α is a common root of f and f', then $f = (x-\alpha)q$ gives $f' = q + (x-\alpha)q'$ and $f'(\alpha) = q(\alpha) = 0$, so $(x-\alpha)^2 \mid f$.

Gcd is the same in $F[x]$ and $\overline{F}[x]$ —Euclidean algorithm stays inside the base ring.

Application: $x^{p^n} - x$ Has p^n Distinct Roots

Take $f = x^{p^n} - x \in \mathbb{F}_p[x]$.

- $f'(x) = p^n\, x^{p^n-1} - 1 = -1$ in characteristic p (every multiple of p vanishes).
- $\gcd(f, f') = \gcd(f, -1) = 1$ —so no repeated roots.
- $\overline{\mathbb{F}_p}$ algebraically closed $\Rightarrow$ f splits into $\deg f = p^n$ linear factors.
- No repeats $\Rightarrow$ p^n *distinct* roots: $\#\{a \in \overline{\mathbb{F}_p} : a^{p^n} = a\} = p^n$.

Existence Theorem

Theorem. $\mathbb{F}_{p^n}$ is a subfield of $\overline{\mathbb{F}_p}$ containing $\mathbb{F}_p$, with exactly p^n elements.

Proof, in two lines.

- Subfield: by the fixed-subfield theorem applied to $\langle \sigma^n \rangle$; contains $\mathbb{F}_p$ by Fermat.
- Size: $x^{p^n} - x$ has $\deg = p^n$ and $\gcd$ with $f' = -1$ is 1, so all p^n roots are distinct in the algebraic closure. $\square$

Uniqueness Theorem

Theorem. Any finite field F with $|F| = p^n$ is $\mathbb{F}_p$-isomorphic to $\mathbb{F}_{p^n}$.

Proof sketch.

- $F/\mathbb{F}_p$ algebraic of degree n; embed $\iota : F \hookrightarrow \overline{\mathbb{F}_p}$ by the universal property of the algebraic closure.
- $|F^\times| = p^n - 1$, so $a^{p^n} = a$ for every $a \in F$ (Lagrange). Hence $\iota(a) \in \mathbb{F}_{p^n}$.
- Counts match ($|\iota(F)| = p^n = |\mathbb{F}_{p^n}|$), so ι is surjective onto $\mathbb{F}_{p^n}$. □

Subfield Lattice: $\mathbb{F}_{p^d} \subseteq \mathbb{F}_{p^n} \iff d \mid n$

Theorem. Inside $\overline{\mathbb{F}_p}$:

$$\mathbb{F}_{p^d} \subseteq \mathbb{F}_{p^n} \quad \iff \quad d \mid n.$$

Proof sketch.

- ($\Leftarrow$) $n = dk$: iterate $a^{p^d} = a$ to get $a^{p^n} = a$.
- ($\Rightarrow$) $\mathbb{F}_{p^n}$ is a vector space over $\mathbb{F}_{p^d}$ of some dim k, so $p^n = (p^d)^k$, giving $n = dk$.

□

Example. $\mathbb{F}_{p^{12}}$ has subfields exactly at $d \in \{1, 2, 3, 4, 6, 12\}$ (divisors of 12).

$\mathbb{F}_{p^n}^{\times}$ Is Cyclic (Fraleigh 33.4)

Theorem. Every finite subgroup G of $F^{\times}$ is cyclic. In particular $\mathbb{F}_{p^n}^{\times}$ is cyclic of order $p^n - 1$.

Proof via the classification of finite abelian groups. G decomposes as a direct product of its Sylow subgroups (using ℓ for the Sylow prime to avoid clashing with the characteristic): $G = \prod_{\ell} G_{\ell}$. Each G_{ℓ} is a finite abelian ℓ-group, so

$$G_{\ell} \;\cong\; \mathbb{Z}/\ell^{a_1} \times \cdots \times \mathbb{Z}/\ell^{a_k}.$$

Bound the number of ℓ-torsion elements. In this product, the elements x with $x^{\ell} = e$ form a subgroup of order ℓ^k (each $\mathbb{Z}/\ell^{a_i}$ contributes its unique subgroup of order ℓ). But $x^{\ell} - 1$ has at most ℓ roots in F, so $\ell^k \leq \ell$, hence $k \leq 1$.

So every Sylow G_{ℓ} is cyclic. G is a product of cyclic groups of pairwise coprime orders, hence cyclic. $\square$

Primitive Element of $\mathbb{F}_{p^n}$

Since $\mathbb{F}_{p^n}^{\times}$ is cyclic, pick a generator α.

- Every nonzero element of $\mathbb{F}_{p^n}$ is a power of α.
- $\mathbb{F}_{p^n} = \mathbb{F}_p(\alpha)$ — a simple extension.
- The minimal polynomial $m_\alpha \in \mathbb{F}_p[x]$ is irreducible of degree n.

Corollary. $\mathbb{F}_p[x]$ contains an irreducible polynomial of every degree $n \geq 1$.

This is the *primitive element theorem* for finite fields.

$\mathrm{Aut}(\mathbb{F}_{p^n}/\mathbb{F}_p) = \langle\sigma\rangle$

Theorem. $\mathrm{Aut}(\mathbb{F}_{p^n}/\mathbb{F}_p) = \langle\sigma\rangle \cong \mathbb{Z}/n\mathbb{Z}$, generated by Frobenius.

Proof.

- σ fixes $\mathbb{F}_p$ (Fermat) and preserves $\mathbb{F}_{p^n}$, so σ is an $\mathbb{F}_p$-automorphism.
- $\sigma^n = \mathrm{id}$ on $\mathbb{F}_{p^n}$ by definition; $\sigma^k \neq \mathrm{id}$ for $0 < k < n$ (else $\mathbb{F}_{p^n} \subseteq \mathbb{F}_{p^k}$, impossible by size). So $|\langle\sigma\rangle| = n$.
- Take a primitive element α with $\mathbb{F}_{p^n} = \mathbb{F}_p(\alpha)$ and $\deg m_\alpha = n$. Any $\mathbb{F}_p$-automorphism τ is determined by $\tau(\alpha)$, and $\tau(\alpha)$ must be a root of m_α. Since m_α has at most n roots, $|\mathrm{Aut}(\mathbb{F}_{p^n}/\mathbb{F}_p)| \leq n$.
- Combined with $|\langle\sigma\rangle| = n$ and $\langle\sigma\rangle \subseteq \mathrm{Aut}$: equality, so $\mathrm{Aut} = \langle\sigma\rangle$. □

$|\mathrm{Aut}(\mathbb{F}_{p^n}/\mathbb{F}_p)| = [\mathbb{F}_{p^n} : \mathbb{F}_p]$. An extension satisfying this equality is called *Galois* — so every $\mathbb{F}_{p^n}/\mathbb{F}_p$ is Galois.

Examples: $\mathbb{F}_4$, $\mathbb{F}_8$, $\mathbb{F}_9$

$\mathbb{F}_4 = \mathbb{F}_2[x]/(x^2+x+1)$. x^2+x+1 has no roots in $\mathbb{F}_2$, irreducible. $\alpha^2 = \alpha + 1$; elements $\{0, 1, \alpha, \alpha+1\}$; $\alpha^3 = 1$, so $\mathbb{F}_4^\times \cong \mathbb{Z}/3$.

$\mathbb{F}_8 = \mathbb{F}_2[x]/(x^3+x+1)$. Cubic, no roots in $\mathbb{F}_2$, hence irreducible. $\mathbb{F}_8^\times \cong \mathbb{Z}/7$ (prime order): every element other than 1 is a generator.

$\mathbb{F}_9 = \mathbb{F}_3[x]/(x^2+1)$. Squares in $\mathbb{F}_3$ are $\{0, 1\}$, so -1 is not a square; x^2+1 is irreducible. With $i^2 = -1 = 2$, elements $\{a + bi : a, b \in \mathbb{F}_3\}$ — the characteristic-3 analogue of $\mathbb{R}(i) = \mathbb{C}$.

Application: Reed–Solomon Error-Correcting Codes

Used in QR codes (over $\mathbb{F}_{2^8} = \mathbb{F}_{256}$), CDs, DVDs, satellite and deep-space communication, RAID storage.

Setup. Fix $\mathbb{F}_q$ and pick n distinct points $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q$ (typically $n \le q$; often α_i are powers of a primitive root).

Encoding. A message $(m_0, \ldots, m_{k-1}) \in \mathbb{F}_q^k$ becomes the polynomial $p(x) = m_0 + m_1 x + \cdots + m_{k-1} x^{k-1}$ of degree $< k$, transmitted as the evaluation vector

$$(p(\alpha_1), p(\alpha_2), \ldots, p(\alpha_n)) \in \mathbb{F}_q^n.$$

Why it corrects errors. Two distinct messages give $p \neq p'$ with $\deg(p - p') < k$, so $p - p'$ has at most $k - 1$ roots (root bound). The two codewords therefore agree in at most $k - 1$ positions — i.e. they differ in at least $d := n - k + 1$ positions (minimum distance).

If the channel injects $\le \lfloor (d-1)/2 \rfloor = \lfloor (n-k)/2 \rfloor$ errors, the received word lies strictly closer (in Hamming distance) to the true codeword than to any other, so decoding is unique.

Reed–Solomon: Worked Example over $\mathbb{F}_7$

Parameters. $q = 7$, $k = 3$, $n = 5$; evaluation points $1, 2, 3, 4, 5 \in \mathbb{F}_7$. Tolerates $\lfloor (5-3)/2 \rfloor = 1$ error.

Encoding. Message $(3, 1, 2) \mapsto p(x) = 3 + x + 2x^2$. Compute (mod 7):

$$p(1) = 6, \quad p(2) = 13 \equiv 6, \quad p(3) = 24 \equiv 3, \quad p(4) = 39 \equiv 4, \quad p(5) = 58 \equiv 2.$$

Transmit codeword $(6, 6, 3, 4, 2)$.

Channel error. Receive $(6, 6, 3, \mathbf{0}, 2)$ (position 4 corrupted).

Decoding. Find the unique p of degree < 3 agreeing with the received word in $\geq n - 1 = 4$ positions.

- Interpolate $(1, 6), (2, 6), (3, 3)$ by Lagrange: $\tilde{p}(x) = 3 + x + 2x^2$.
- Check remaining: $\tilde{p}(4) = 4 \neq 0$ (mismatch), $\tilde{p}(5) = 2$ (match).
- Only 1 mismatch $\leq$ tolerated errors — accept. Decoded message: $(3, 1, 2)$, error at position 4.

Practical Decoding: One Linear System (Welch–Berlekamp)

Trying every k-subset is $\binom{n}{k}$ work — exponential. In practice: decoding is a single linear system in $\mathbb{F}_q$.

Setup. Let $E(x) = \prod_{i \text{ error pos}}(x - \alpha_i)$ be the (unknown) *error-locator polynomial*: monic, degree = number of errors $\leq e := \lfloor (n-k)/2 \rfloor$. Define $Q(x) := p(x)\,E(x)$ where p is the original message polynomial. Then $\deg Q < k + e$.

Key identity. For every received position i:

$$y_i \cdot E(\alpha_i) \;=\; Q(\alpha_i).$$

(Non-error: $y_i = p(\alpha_i)$, multiply by $E(\alpha_i)$. Error: $E(\alpha_i) = 0$, so both sides vanish.)

Algorithm. Solve n linear equations $y_i E(\alpha_i) = Q(\alpha_i)$ for the unknown coefficients of E (degree $\leq e$, monic) and Q (degree $< k + e$) — $2e + k$ unknowns. Then $p = Q/E$ (polynomial division); roots of E give error positions.

Complexity. $O(n^3)$ Gaussian elimination, $O(n^2)$ via Berlekamp–Massey (a specialized linear-recurrence solver that exploits the system's structure), $O(n \log^2 n)$ with FFT. QR code ($n = 256$): microseconds.